

Llemma: An Open Language Model For Mathematics

Year: 2023 | Citations: 370 | Authors: *Zhangir Azerbayev, Hailey Schoelkopf, Keiran Paster*

Abstract

We present Llemma, a large language model for mathematics. We continue pretraining Code Llama on the Proof-Pile-2, a mixture of scientific papers, web data containing mathematics, and mathematical code, yielding Llemma. On the MATH benchmark Llemma outperforms all known open base models, as well as the unreleased Minerva model suite on an equi-parameter basis. Moreover, Llemma is capable of tool use and formal theorem proving without any further finetuning. We openly release all artifacts, including 7 billion and 34 billion parameter models, the Proof-Pile-2, and code to replicate our experiments.